

CLAIMS

We claim:

1 1. A method of integrating a document in a first format into a data store holding
2 documents in a second format, the method comprising:
3 supplying the document in the first format and a specification comprising
4 instructions for creating a description of the document based on
5 attributes of the document and syntax rules for the description;
6 receiving the document in the second format;
7 receiving a description of the document generated responsive to the
8 specification; and
9 importing the document in the second format into the data store responsive to
10 the description.

1 2. The method of claim 1 wherein the attributes disclosed of the document
2 include at least one of the creation date of the document, the source of the document,
3 content contained in the document and the location of the document on a storage
4 medium.

1 3. The method of claim 1 further comprising:
2 receiving the document in the second format and the description of the
3 document as part of a batch file also containing a plurality of other
4 documents in the second format and associated descriptions of the
5 plurality of other documents; wherein the other documents in the
6 second format are configured to be imported into the data store

7 responsive to the associated descriptions of the other documents in the
8 second format.

1 4. The method of claim 1 further comprising indexing the document imported
2 into the data store based on indexing data contained in the description.

1 5. The method of claim 1 wherein the specification comprises an XML
2 Document Type Definition that describes element names and XML syntax rules for
3 creating a description of the document.

1 6. The method of claim 6 wherein the description comprises a well-formed XML
2 document file generated responsive to the XML Document Type Definition.

1 7. The method of claim 1, wherein the document in the first format comprises a
2 paper document, and the document in the second format comprises an electronic file.

1 8. A system for integrating a plurality of documents in a first format into a data
2 store holding documents in a second format, the system comprising:
3 a repository configured to store a plurality of documents in the first format
4 and a specification comprising instructions for creating descriptions of
5 the plurality of documents based on attributes of the documents and
6 syntax rules, the repository further configured to supply the documents
7 and specification to a conversion facility;
8 a batch import module adapted to receive from the conversion facility the
9 plurality of documents in the second format and descriptions of the

10 plurality of documents in the second format generated responsive to
11 the specification, wherein the batch import module is further adapted
12 to import the plurality of documents in the second format responsive to
13 the descriptions into the data store; and wherein the data store is
14 further configured to provide access to a user to the plurality of
15 documents in the second format.

1 9. The system of claim 8 wherein a single batch file contains the plurality of
2 documents in the second format and the descriptions of the plurality of documents in
3 the second format, and the batch import module receives the plurality of documents in
4 the second format and the descriptions of the plurality of documents in the second
5 format in the form of the single batch file.

1 10. The system of claim 8 wherein the descriptions contain indexing data and the
2 data store is further adapted to store references in an index to the plurality of documents
3 imported into the data store responsive to the indexing data contained in the
4 descriptions.

1 11. The system of claim 8 wherein the specification comprises an XML
2 Document Type Definition that describes element names and XML syntax rules for
3 creating a description of the document.

1 12. The system of claim 11 wherein the description comprises a well-formed
2 XML document file generated responsive to the XML Document Type Definition

1 13. The system of claim 8, wherein the plurality of documents in the first format
2 comprise paper documents, and the plurality of documents in the second format
3 comprise electronic files.

1 14. The system of claim 8, wherein the specification further comprises:
2 instructions for storing documents with shared attributes in a common batch
3 file, creating a batch file default description of the documents with
4 shared attributes responsive to the shared attributes of the documents,
5 and using the batch default description to create descriptions of the
6 documents with shared attributes.

1 15. A computer-implemented method for integrating electronic files into a data
2 store responsive to descriptions of the files, the method comprising:
3 receiving the electronic files and the descriptions of the files, the descriptions
4 descriptive of attributes of the electronic files and generated
5 responsive to a specification comprising instructions for describing the
6 files and syntax rules for the descriptions;
7 locating the electronic files on a storage medium based on location
8 information contained within the descriptions;
9 copying the electronic files into the data store;
10 extracting indexing data associated with the electronic files from the
11 descriptions of the electronic files; and
12 indexing the electronic files in the data store responsive to the indexing data
13 extracted from the descriptions of the electronic files.

1 16. The method of claim 15 further comprising creating references in an index to
2 the electronic files in the data store responsive to the indexing data to enable
3 subsequent access to the files by a user application using the index.

1 17. The method of claim 15 wherein the electronic files and the descriptions of
2 the files are stored in a single batch and further comprising:
3 receiving the electronic files and the descriptions of the files in the form of the
4 single batch.

1 18. The method of claim 15 further comprising indexing the electronic files in the
2 data store responsive to batch-level indexing data extracted from the descriptions of the
3 electronic files.

1 19. The method of claim 15 wherein the step of extracting indexing data about the
2 electronic files from the descriptions of the electronic files is performed by a parser and
3 further comprising the steps of:

4 locating valid indexing data about the electronic files contained in the
5 descriptions responsive to the syntax rules in the specification;
6 extracting valid indexing data from the descriptions; and
7 outputting the valid indexing data to the data store.

1 20. A computer implemented batch import apparatus for integrating a plurality of
2 electronic files into a data store, the apparatus comprising:

3 a repository configured to receive the electronic files and the descriptions of

4 the files; the descriptions generated responsive to a specification
5 comprising instructions for describing attributes of the files and syntax
6 rules for the descriptions;
7 a file import module adapted to locate the electronic files based on location
8 information contained within the descriptions of the files and import
9 the electronic files into the data store; and
10 an indexing module adapted to index the electronic files in the data store
11 responsive to the indexing data extracted from the descriptions of the
12 electronic files.

1 21. The apparatus of claim 20 further comprising a user application module
2 configured to access an electronic file in the data store.

1 22. The apparatus of claim 20 wherein a single batch file contains the electronic
2 files and the descriptions of the files, and the repository receives the electronic files and
3 the descriptions of the files in the form of the single batch file.

1 23. The apparatus of claim 20 further wherein the indexing module indexes the
2 electronic files in the data store responsive to batch-level indexing data extracted from
3 the descriptions of the electronic files.

1 24. The apparatus of claim 20 further comprising a parser for locating valid
2 indexing data about the electronic files contained in the descriptions responsive to the
3 syntax definitions in the specification, extracting valid indexing data from the
4 descriptions, and outputting the valid indexing data to the indexing module.

1 25. The apparatus of claim 20, wherein the specification further comprises
2 instructions for storing files with shared attributes in a common batch, creating a batch
3 default description of the files with shared attributes responsive to the shared attributes
4 of the files, and using the batch default description to create descriptions of the files
5 with shared attributes.

1 26. A computer program product comprising:
2 a computer readable medium; and
3 computer program instructions, encoded on the medium, for controlling a
4 processor to perform the operations of:
5 receiving a document in a second format converted from a document in a first
6 format;
7 receiving a well-formed XML source file describing the document generated
8 responsive to an XML Document Type Definition and descriptive of
9 an attribute of the document;
10 importing the document into a data store responsive to attribute descriptions
11 contained in the XML source file; and
12 accessing the document in the data store.

1 27. A computer program product comprising:
2 a computer readable medium; and
3 computer program instructions, encoded on the medium, for controlling a
4 processor to perform the operations of:
5 receiving electronic files and descriptions of the files, the descriptions

6 descriptive of attributes of the electronic files and generated
7 responsive to a specification comprising instructions for describing the
8 files and syntax rules for the descriptions ;
9 locating the electronic files on a storage medium based on location
10 information contained within the descriptions ;
11 copying the electronic files into a data store;
12 extracting indexing data about the electronic files from the descriptions of the
13 electronic files; and
14 indexing the electronic files in the data store responsive to the indexing data
15 extracted from the descriptions of the electronic files.

1 28. The computer program product of claim 27, further comprising:
2 computer program instructions, encoded on the medium, for controlling a
3 processor to perform the operation of:
4 creating references in an index to the electronic files in the data store
5 responsive to the indexing data to enable subsequent access to the files
6 by a user application using the index.

1 29. The computer program product of claim 27, further comprising:
2 computer program instructions, encoded on the medium, for controlling a
3 processor to perform the operation of:
4 indexing the electronic files in the data store responsive to batch-level
5 indexing data extracted from the descriptions of the electronic files.

1 30. The computer program product of claim 27, further comprising:
2 computer program instructions, encoded on the medium, for controlling a
3 processor to perform the operations of:
4 locating valid indexing data about the electronic files contained in the
5 descriptions responsive to the syntax rules in the specification;
6 extracting valid indexing data from the descriptions; and
7 outputting the valid indexing data to the data store.